



GENOME ANNOUNCEMENT

Open Access

Comparing the genomes of *Helicobacter pylori* clinical strain UM032 and Mice-adapted derivatives

Yalda Khosravi¹, Vellaya Rehvathy¹, Wei Yee Wee², Susana Wang³, Primo Baybayan³, Siddarth Singh⁴, Meredith Ashby³, Junxian Ong⁵, Arlaine Anne Amoyo⁵, Shih Wee Seow⁵, Siew Woh Choo², Tim Perkins⁶, Eng Guan Chua⁷, Alfred Tay⁷, Barry James Marshall⁷, Mun Fai Loke¹, Khean Lee Goh⁸, Sven Pettersson^{5,9,10} and Jamuna Vadivelu^{1*}

Abstract

Background: *Helicobacter pylori* is a Gram-negative bacterium that persistently infects the human stomach inducing chronic inflammation. The exact mechanisms of pathogenesis are still not completely understood. Although not a natural host for *H. pylori*, mouse infection models play an important role in establishing the immunology and pathogenicity of *H. pylori*. In this study, for the first time, the genome sequences of clinical *H. pylori* strain UM032 and mice-adapted derivatives, 298 and 299, were sequenced using the PacBio Single Molecule, Real-Time (SMRT) technology.

Result: Here, we described the single contig which was achieved for UM032 (1,599,441 bp), 298 (1,604,216 bp) and 299 (1,601,149 bp). Preliminary analysis suggested that methylation of *H. pylori* genome through its restriction modification system may be determinative of its host specificity and adaptation.

Conclusion: Availability of these genomic sequences will aid in enhancing our current level of understanding the host specificity of *H. pylori*.

Keywords: *Helicobacter pylori*, PacBio Single Molecule, Real-Time (SMRT) technology, Clinical *H. pylori*, Mice-adapted

Background

Helicobacter pylori persistently colonizes the human stomach to cause chronic gastritis, peptic ulcer disease, gastric adenocarcinoma, and gastric mucosa-associated lymphoid tissue (MALT) lymphoma [1]. The mechanisms involved in the pathogenesis of *H. pylori* infections are still not fully established [2]. Thus, experimental animal models that mimic human diseases are essential to provide information on etiopathogeny, immunity and therapy, as well as to improve our understanding on ways *H. pylori* can induce a diverse range of gastric pathologies [3,4]. Among various animal models available, mouse remains the most readily used animal model for studying *H. pylori*-induced diseases and have played important roles in the elucidation of

factors required for colonization, distribution and persistence of infection [2].

Methods

Mice adaptation study

Adopting a similar strategy as in previous studies [2,5], a pool consisting of twelve clinical strains of *H. pylori* isolated from patients presenting for gastroscopy at the University of Malaya Medical Centre (UMMC) was inoculated intragastrically into five 4–6 weeks old male C57BL/6 mice. Multiple colonies of *H. pylori* were successfully recovered from the gastric tissue sample of a mouse (1/5) following necropsy two weeks post-infection. Random amplification of polymorphic DNA (RAPD) fingerprinting was used to trace back the mice-adapted isolates to its parental clinical strain, UM032. *H. pylori* UM032 was isolated from a patient presenting with peptic ulcer disease. Mice-adapted

* Correspondence: jamuna@um.edu.my

¹Department of Medical Microbiology, University of Malaya, Kuala Lumpur, Malaysia

Full list of author information is available at the end of the article

isolates from the first mouse passage were designated as 298 and were used for the second round of mouse passage to access the stability and infectivity of this mice-adapted strain. All three mice inoculated with the mice-adapted 298 strain were successfully infected and *H. pylori* isolated from the second passage were designed as 299. The animal study was performed with the approval of the SingHealth Institutional Research Committees (SHS IBC) and the Ethical Committee for Animal Research (Form No. SHS-IBC-201, January 2010).

Genome sequencing

In this study, *H. pylori* DNA was isolated using the RTP Bacteria DNA Mini Kit (Invitex GmbH, Berlin, Germany). The extracted DNA samples were sequenced using Pacific Biosciences RS sequencing technology (Pacific Biosciences, Menlo Park, CA), yielding >20× average genome coverage. Each sample was prepared as a 10-kb insert library using C2 chemistry and sequenced on 8 Single-Molecule Real-Time (SMRT) cells.

Assembly and annotation

De novo assembly of the read sequences was created using the continuous long reads (CLR) following the Hierarchical Genome Assembly Process (HGAP) workflow (<http://pacbiodevnet.com/>) as available in SMRT Analysis v2.0. The genomes were annotated with the NCBI (National Center for Biotechnology Information) Prokaryotic Genomes Automatic Annotation Pipeline and NMPDR (National Microbial Data Resource) Rapid Annotation using Subsystem Technology (RAST) [6].

The SEED-Viewer was used to visualize the genome annotation and comparison generated by RAST [7].

Submission of genome sequence

The genome sequence of the *Helicobacter pylori* strains UM032, 298 and 299 are available in DDBJ/EMBL/GenBank under Accession numbers CP005490, CP006610 and CP005491 respectively.

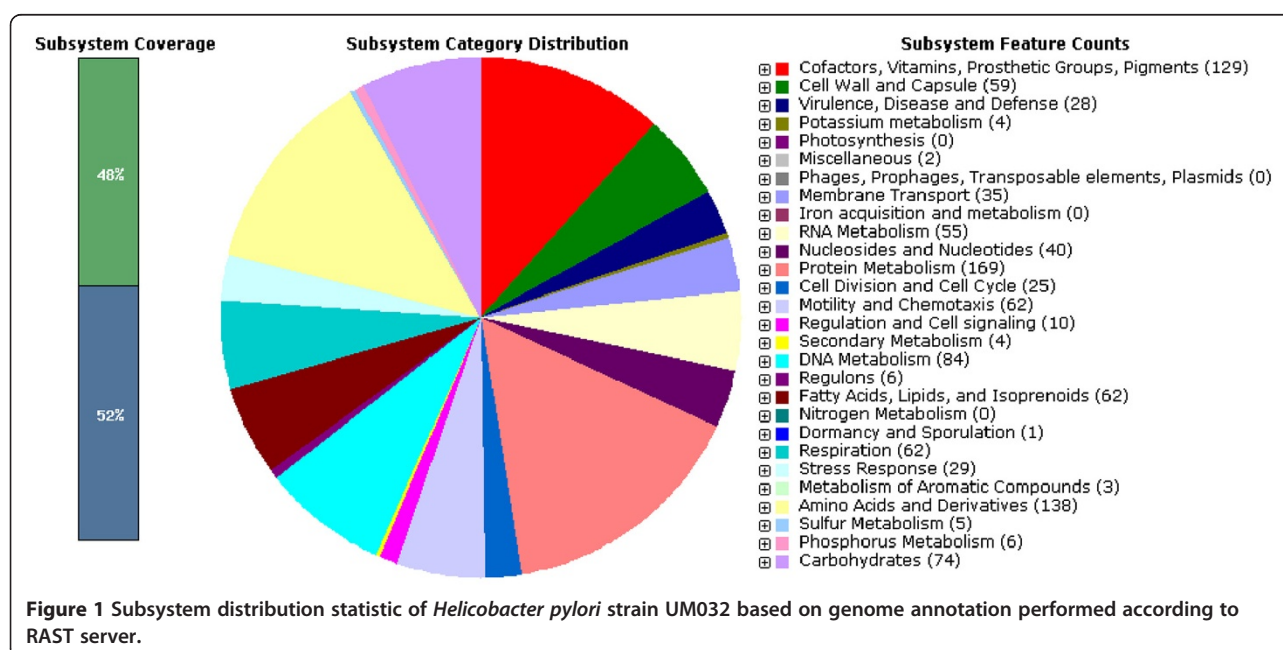
Quality assurance

The genomic DNA was isolated from pure bacterial isolate (positive for urease, catalase and oxidase tests) and was further confirmed with *16S rRNA* sequencing and genotyping of bacterial virulence factors. Bioinformatic assessment of potential contamination of the genomic library by allochthonous microorganisms was done using PGAAP and RAST annotation systems.

Initial findings

Genome characteristics

Based on the assembled genomes with HGAP using PacBio long reads from a single library preparation, single contigs were achieved for UM032 (1,599,441 bp), 298 (1,604,216 bp) and 299 (1,601,149 bp). The GC content for all three assembled genomes was 38.8%. Additional information is included in the sequencing reports: UM032 (Additional file 1), 298 (Additional file 2) and 299 (Additional file 3). Figure 1 describes the subsystem distribution of the parental clinical strain, UM032. Figure 2 shows the sequence homology between 298, 299 and Shi470 with reference to UM032. *H. pylori* Shi470 was predicted to be among those closest to UM032 with



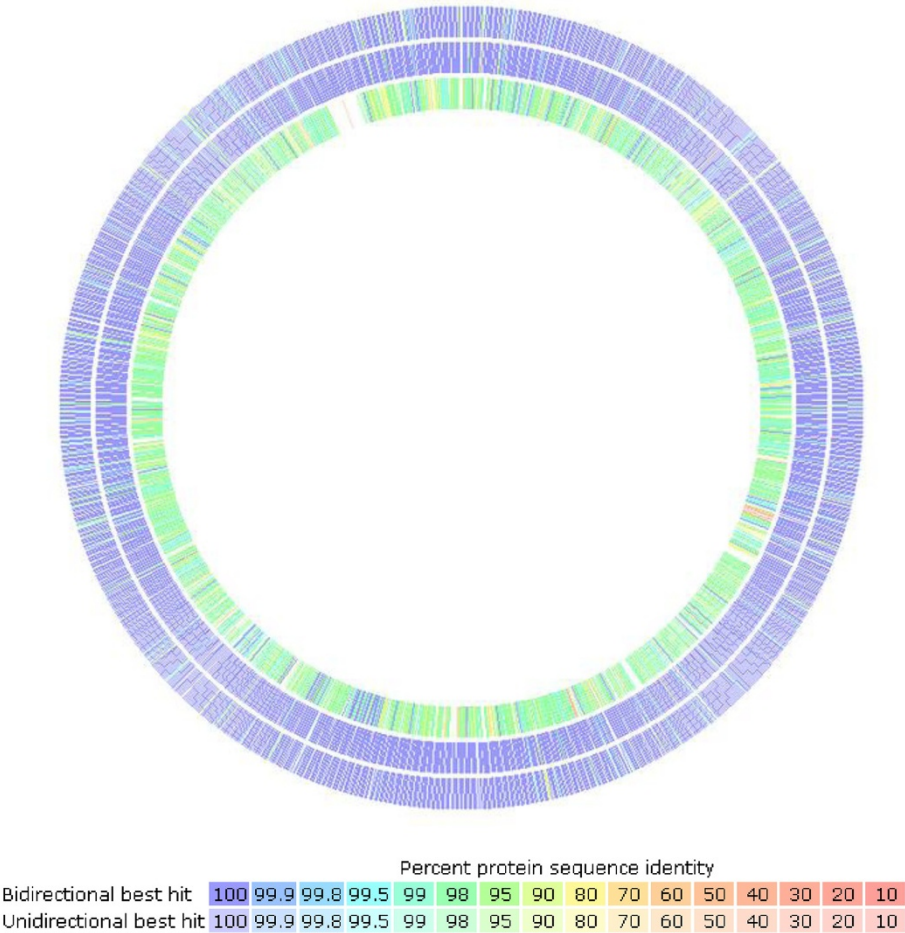


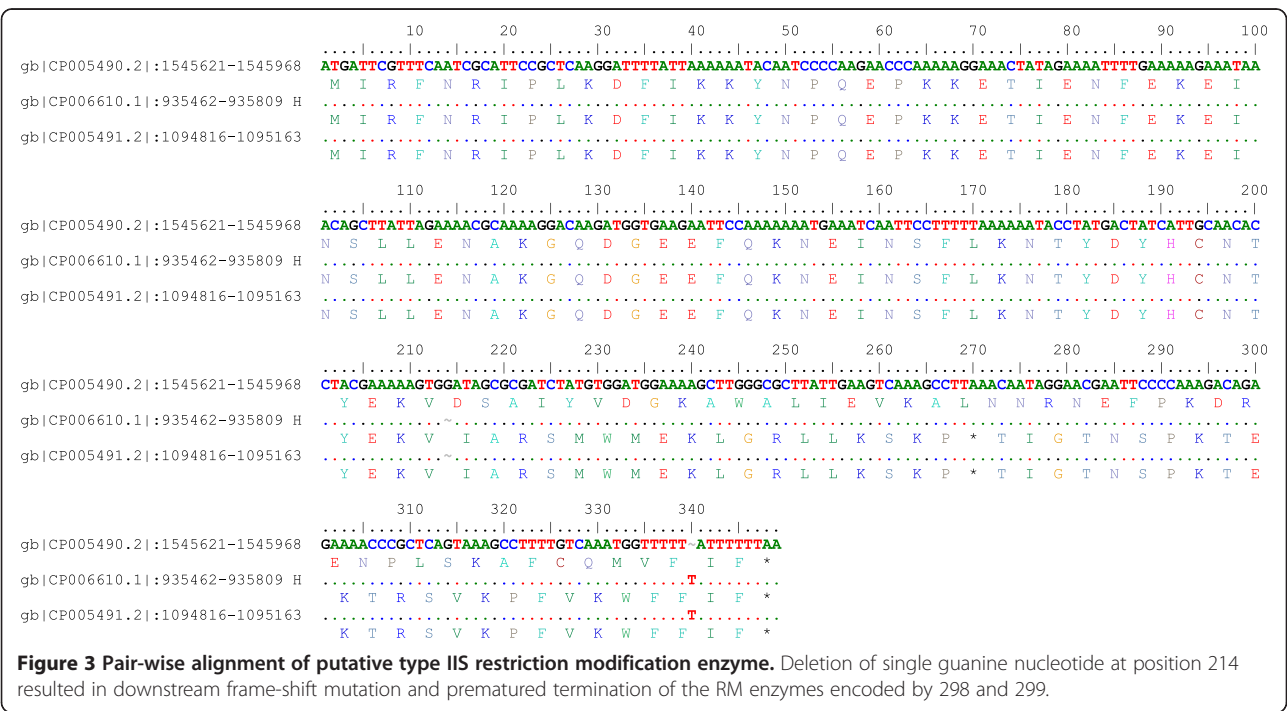
Figure 2 Genome sequence comparison of *Helicobacter pylori* 298 (outer) 299 (middle) and Shi470 (inner) when aligned with reference genome, UM032, using RAST program. Intensity of color indicates degree of protein identity (legend).

score of 470. Neither gene lost nor gain was found in the mice-adapted derivatives (298 and 299) when compared to the parental strain (UM032). Interestingly, total base modifications detected through the PacBio RS sequencing platform as described under Table 1 was reducing with passaging in mice. In addition, a 348 b.p. gene encoding for a putative type IIS restriction modification (R-M) enzyme in *H. pylori* UM032 was found in to truncated in 298 and 299 (Figure 3). Thus, methylation may be a mean of host adaptation by *H. pylori* and may have an important role in determining host specificity.

The availability of complete sequences of mice-adapted strains and their parental clinical isolate will provide important information that contributes towards our understanding of the host specific and adaptation of *H. pylori*. In addition, it will help in extrapolate results obtained using mice model to the natural human host of *H. pylori*. *H. pylori* 298 strain will be used for *H. pylori* colonizing studies in mice.

Table 1 Type of base modifications and associated motifs detected

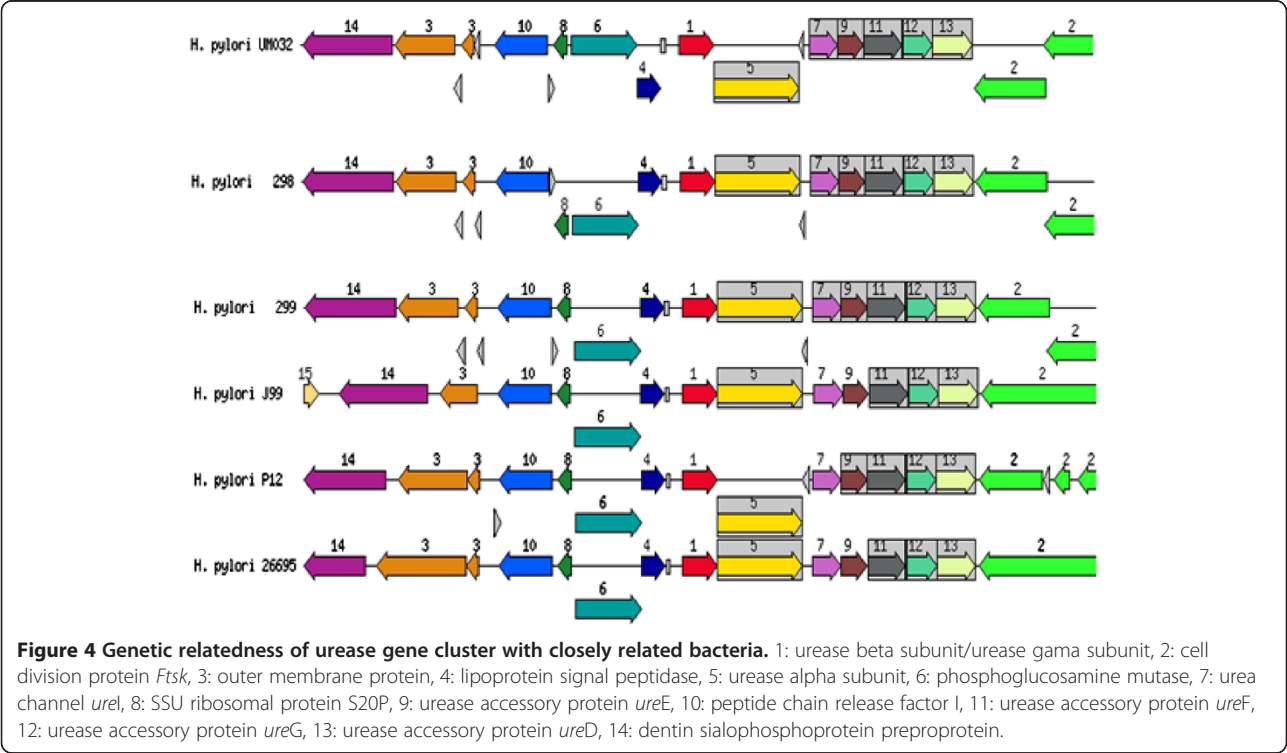
Motif	Modification type	# of motifs detected		
		UM032	298	299
GANTC	m6A	5,393	5,428	5,397
CCATC	m6A	2,257	2,261	2,258
GAGG	m6A	4,585	4,598	4,580
TCNGA	m6A	2,531	2,544	2,534
GATC	m6A	10,195	10,210	10,175
CCGG	m4C	3,414	3,424	3,420
TGCA	m6A	11,221	11,199	11,185
CYANNNNNNTRG	m6A	2,303	2,319	2,305
ATTAAT	m6A	865	865	865
ACNGT	m4C	1,077	1,056	1,056
CATG	m6A	13,446	13,339	13,361
GAAAG	Unknown	4,332	4,839	4,851
Others		59,182	43,338	36,469
Total		120,801	105,420	98,456



Putative gene clusters responsible for survival and virulence of *H. pylori*

H. pylori possess genes for cytosolic urease biosynthesis, which is governed by a seven-gene cluster, are essential for its survival in the acidic gastric environment [8].

H. pylori vacuolating cytotoxin A (*VacA*) is an important virulence factor of the bacterium [9]. Using the SEED database, genetic relatedness of the urease gene cluster and *vacA* for the clinical strain (UM032) and its mice-adapted counterparts (298 and 299) in comparison



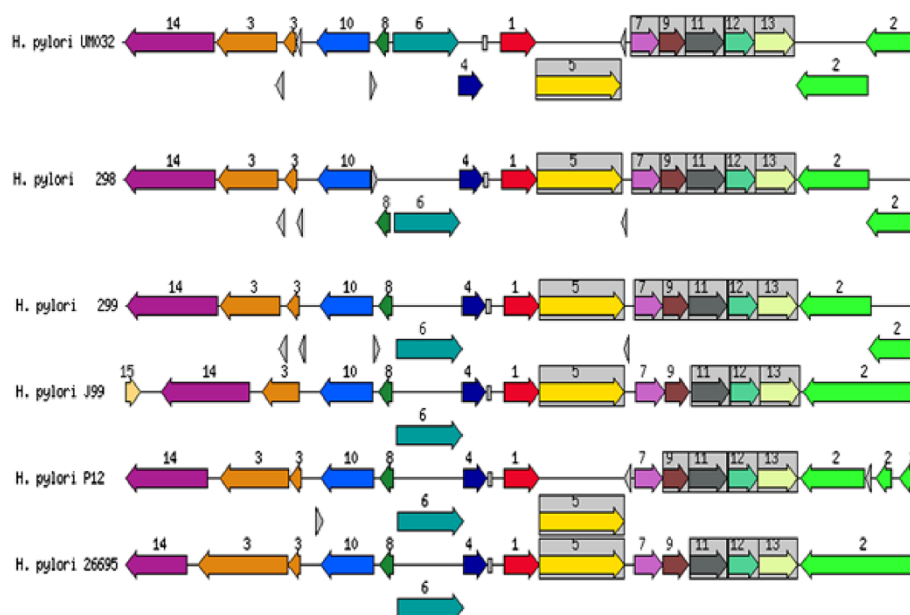


Figure 5 Genetic relatedness of *vacA* cluster with closely related bacteria. 1: vacuolating cytotoxin, 2: hypothetical protein, 3: haem iron uptake system ATP-binding protein, 4: cysteineyl-Trna-SYNTHETASE, 5: IRON III, 6: dehydrogenases with different specificities, 7: proposed peptidoglycan lipid, 8: hypothetical protein, 9: hypothetical protein, 10: DNA damage inducible protein J, 11: holliday junction DNA helicase RUUA, 12: putative outer membrane protein, 13: hypothetical protein.

to other known *H. pylori* strains are shown in Figures 4 and 5.

Future directions

To our knowledge, this is the first genome sequence of *H. pylori* isolated from human and mouse using PacBio SMRT Technology. Comparative genomic and more-detailed methylomic analysis of these data is in process and will be included in future publications. Mice-adapted *H. pylori* described here will be used in future *H. pylori* infection studies in mice.

Availability of supporting data

The data sets supporting the results of this article are included within the additional files.

Additional files

Additional file 1: Assembly report for UM032.

Additional file 2: Assembly report for 298.

Additional file 3: Assembly report for 299.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

YK, MFL, JV, SSW and SP planned the experiments. YK and MFL prepared the *H. pylori* for inoculation, isolated the bacteria from mice, performed genetic fingerprinting and extracted the DNA. JO, AAA and SWS were responsible for animal infection and study. SW, PB, SS and MA carried out the genomic sequencing and initial bioinformatics. VR, TP, ENG, AT carried out further

bioinformatics analysis. BJM, JV, KLG, SP and MFL are senior co-authors. All authors have read the manuscript and approved.

Acknowledgements

Funding was provided by the University of Malaya-Ministry of Higher Education (UM-MOHE) High Impact Research (HIR) grant (reference UM.C/625/1/HIR/MOHE/CHAN-02; "Molecular Genetics"). SW, PB, SS and MA are full-time employees at Pacific Biosciences, the commercial company for the SMRT sequencing technologies. We thank the NCBI (National Center for Biotechnology Information) rapid annotation pipeline team for providing the genome annotation services.

Author details

¹Department of Medical Microbiology, University of Malaya, Kuala Lumpur, Malaysia. ²Dental Research and Training Unit, Faculty of Dentistry, University of Malaya, Kuala Lumpur, Malaysia. ³Pacific Biosciences, Menlo Park, California, USA. ⁴PacBio Singapore, Singapore, Singapore. ⁵National Cancer Centre, Singapore, Singapore. ⁶School of Pathology and Laboratory Medicine, Faculty of Medicine, Dentistry and Health Sciences, University of Western Australia, Perth, Western Australia, Australia. ⁷The Marshall Centre for Infectious Diseases Research and Training, University of Western Australia, Perth, Western Australia, Australia. ⁸Department of Medicine, University of Malaya, Kuala Lumpur, Malaysia. ⁹Department of Microbiology, Tumor and Cell Biology (MTC), Karolinska Institutet, Stockholm, Sweden. ¹⁰School of Biological Sciences, Nanyang Technological University, Singapore, Singapore.

Received: 2 July 2013 Accepted: 16 August 2013

Published: 19 August 2013

References

1. Kusters JG, van-Vliet AHM, Kuipers EJ: **Pathogenesis of *Helicobacter pylori* Infection.** *Clin Microbiol Rev* 2006, **19**:449–490.
2. Fox J, Lee A: **The role of *Helicobacter* species in newly recognized gastrointestinal tract diseases of animals.** *Lab Anim Sci* 1997, **47**:222–255.
3. Lee A, O'Rourke J, Ungria M, Robertson B, Daskalopoulos G, Dixon M: **A standardized mouse model of *Helicobacter pylori* infection: introducing the Sydney Strain.** *Gastroenterol* 1997, **112**:1386–1397.

4. Salama NR, Otto G, Tompkins L, Falkow S: **Vacuolating cytotoxin of *Helicobacter pylori* plays a role during colonization in a mouse model of infection.** *Infect Immun* 2001, **69**:730–736.
5. O'Rourke JL, Lee A: **Animal models of *Helicobacter pylori* infection and disease.** *Microbes Infect* 2003, **8**:741–748.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisano K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O: **The RAST Server: rapid annotations using subsystems technology.** *BMC Genomics* 2008, **8**:9–75.
7. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de-Cr  cy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V et al: **The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes.** *Nucleic Acids Res* 2005, **33**:5691–5702.
8. Kulkarni G, Dhotre D, Dharne M, Shetty S, Chowdhury S, Misra V, Misra S, Patole M, Shouche Y: **Draft genome of *Ochrobactrum intermedium* strain M86 isolated from non-ulcer dyspeptic individual from India.** *Gut Pathog* 2013, **5**:7. doi:10.1186/1757-4749-5-7.
9. Atherton JC, Cao P, Peek RM Jr, Tummuru MK, Blaser MJ, Cover TL: **Mosaicism in vacuolating cytotoxin alleles of *Helicobacter pylori*. Association of specific vacA types with cytotoxin production and peptic ulceration.** *J Biol Chem* 1995, **270**:17771–17777.

doi:10.1186/1757-4749-5-25

Cite this article as: Khosravi et al.: Comparing the genomes of *Helicobacter pylori* clinical strain UM032 and Mice-adapted derivatives. *Gut Pathogens* 2013 **5**:25.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

